

Report from the

# **Employment Rate Measurement Methodology Work Group (Committee on Statistics)**

TO THE  
PRESIDENTIAL  
TASK FORCE  
ON EMPLOYMENT  
OF ADULTS WITH  
DISABILITIES

*Chair:*

*Philip Rones*, Assistant Commissioner  
Bureau of Labor Statistics  
Office of Employment and Unemployment Statistics  
Department of Labor

U.S. Department of Labor

Bureau of Labor Statistics  
2 Massachusetts Ave. N.E.  
Washington, D.C. 20212



The Honorable Alexis Herman  
Chairwoman, Presidential Task Force on  
Employment of Adults with Disabilities

The Honorable Tony Coelho  
Vice-Chair, Presidential Task Force on  
Employment of Adults with Disabilities  
200 Constitution Avenue, N.W., Room S-2312  
Washington, D.C. 20210

Dear Chairwoman Herman & Vice-Chair Coelho:

On behalf of the Employment Rate Measurement Methodology (ERMM) Work Group of the Presidential Task Force on the Employment of Adults with Disabilities (PTFEAD), I am honored to forward to you the PTFEAD Employment Rate Measurement Methodology Work Group Report.

The ERMM Work Group has gathered representatives from nearly 20 Federal agencies with the mandate to design and implement a statistically valid and reliable measure of employment for adults with disabilities. By bringing together the expertise of these dedicated professionals, the Work Group has made significant strides toward achieving its goal.

I would like to thank the Task Force for its invaluable assistance in supporting the work of the ERMM Work Group, and I look forward to continuing this important work to its successful completion.

Sincerely yours,

A handwritten signature in cursive script, reading "Philip L. Rones", is positioned above the typed name.

PHILIP L. RONES  
Chairman  
ERMM Work Group

# The Mandate

*The Bureau of Labor Statistics of the Department of Labor and the Census Bureau of the Department of Commerce, in cooperation with the Departments of Education, Health and Human Services, the National Council on Disability, and the President's Committee on the Employment of People with Disabilities shall design and implement a statistically reliable and accurate method to measure the employment rate of adults with disabilities as soon as possible, but no later than the date of termination of the Task Force. Data derived from this methodology shall be published on as frequent a basis as possible.*

## Background

The development of an employment measure for adults with disabilities has become a critical need for policy makers, analysts, and others concerned with their labor market status. The employment rate, as a measure of labor market activity, is vital to the design, implementation, and evaluation of legislation and programs enabling persons with disabilities to participate as fully as possible in the labor market. Employment data also would tell us how the cyclical expansions and contractions of the economy affect employment among those with disabilities compared to other population groups.

Efforts to produce a statistically accurate and reliable measure of the employment rate for adults with disabilities began over two years prior to the issuance of the Executive Order. In a cooperative initiative undertaken by the Bureau of Labor Statistics (BLS), the Bureau of the Census, the President's Committee on the Employment of People with Disabilities (PCEPD), the Equal Employment Opportunity Commission (EEOC), and the National Council on Disability (NCD), a small number of questions designed to identify this population were drafted and placed in the Survey of Income and Program Participation (SIPP). SIPP also contained a comprehensive module of questions aimed at identifying and characterizing individuals with disabilities. The thought was that if the small set of questions identified the same population as the more extensive

module, then the smaller set could be confidently placed in the Current Population Survey (CPS), the monthly national household survey which collects information used to prepare the official estimates of total employment and unemployment for the Nation and for various population groups. This would allow BLS to produce regular estimates of employment and unemployment for persons with disabilities that would be consistent with the official measures for other groups.

The results of the test were not encouraging. The smaller set of questions failed to identify the same individuals with disabilities who were identified by the full battery of questions. In particular, the vast majority of those identified as having a disability but not a significant disability in the full SIPP disability module were not identified as having a disability with the short set of questions. It later became clear that even the full SIPP battery may have limitations as a "benchmark" for identifying this population. The Bureau of Census found that an extraordinarily high proportion of individuals who indicated they had a disability when first surveyed by SIPP did not respond they had the same disability a year later. One would expect there might be minor fluctuations in disability status over the course of a year but no one expected, for example, that only about 25 percent or fewer individuals who had a significant visual or hearing impairment in the first year would have the same impairment in the second year.

Several possible explanations emerged. One was that the time frame which the questions referred to (long-term) was not clear in the minds of the respondents when they were initially interviewed. However, it may also be that the disability population is very dynamic, and that shifts of persons into and out of different disability status categories are indeed large. Whatever the reasons for the unexpectedly large shifts in disability status indicated by the SIPP data, it is important to understand and explain this phenomenon, if effective questions are to be designed.

The Presidential Task Force on the Employment of Adults with Disabilities came into being at this point. The government agencies that worked with BLS in its initial efforts to develop questions for the CPS also participated in the Employment Rate Measure-

ment Methodology (ERMM) Work Group and the initial measurement effort was subsumed into the work of the ERMM. Currently, approximately 16 government agencies are represented on the Work Group.

Several Work Group meetings have been held to prepare a foundation for comprehensive research that would result in the design of questions to identify persons with disabilities in the context of the CPS. One important discussion involved particular Federal agencies' uses of disability employment data and the dimensions of disability they would like identified through a survey. For example, most agreed that information about the severity of disability is important, and that it would be particularly useful if the survey were able to distinguish employment rates for those with physical and mental disabilities.

The next task identified by the ERMM Work Group was to put together a database containing survey questions on disability along with some measures of the effectiveness of these questions. In order to determine what surveys and survey questions are currently available, and what level of testing the questions have undergone, the BLS compiled an annotated bibliography of survey instruments along with information about the reliability, validity, and other testing the questions have undergone. The bibliography indicated that even the most promising survey questions require additional testing. (The Executive Summary to the Annotated Bibliography is included in the Appendix I.)

## Positive Efforts Already Underway

The Work Group on Employment Rate Measurement Methodology (ERMM) is developing a detailed research plan and a set of research protocols to select or design questions to identify this population. (The research plan is included in the Appendix II.) The first priority of the Work Group is to develop monthly estimates of employment rates for persons with disabilities. Assuming that an effective small set of questions can be developed, and that

this set of questions does not impose an excessive burden on respondents, the goal is to embed it within the monthly CPS. The second priority of the Work Group is to evaluate the relative accuracy of different longer modules of questions in identifying this population.

The Annotated Bibliography (noted above) narrowed the range of surveys and questions that could potentially be used to gather disability data, but it is still difficult to select the surveys in which the Work Group should invest its research efforts.

Conceptually, the research plan is relatively straightforward. Complete batteries of questions on disability will be selected from several surveys that were identified in the Annotated Bibliography as being the most promising from the standpoint of the ERMM Work Group. Several sets of "screener" questions will be identified and/or developed to be tested along with the longer disability assessment surveys (DAS). Screener questions have several important characteristics. They are typically few in number — 1-6 questions, so they minimize the time and space requirements imposed upon an existing survey, such as the CPS. They are designed to identify a population of people with disabilities to whom one can administer a longer survey — one in which specific types of disabilities can be determined.

Then, both the screener and the longer, more detailed complete batteries of questions will be tested on comparable panels of respondents. The data would then be analyzed to determine how well the screener questions identified the same population as the longer survey questions.

The relative accuracy of the different sets of screener questions, however, is only meaningful if the more complete batteries of questions have the same degree of accuracy. The quality of the data generated by the complete batteries of questions may be assessed through various techniques, including comparisons of results from the screener questions and other questions on the CPS, as well as the standard survey evaluation techniques of cognitive interviews, respondent debriefing, and behavior coding.

Ultimately, if the ERMM Work Group research is successful, a set of screener questions and the related complete battery of questions will be selected as the best instruments to identify this population in the context of the CPS. The testing indicated above, however, will largely be conducted outside of the context of the CPS. The number of surveys and questions that must be examined render the CPS an inappropriate research vehicle for this phase of the research. When the list of candidate screener questions is considerably narrowed (to perhaps one or two sets), they can and may be tested in the context in which they will be used — the CPS.

The ERMM Work Group has several options it can explore in the event that none of the screener questions adequately identifies the same disability population as the longer, more detailed surveys.

Even if the screener questions initially fail to identify the same population as the longer disability surveys there is a strong likelihood that the Work Group will have gone far in understanding why the screener questions fail to identify the same populations. We will be able to examine the demographic and disability information in detail to see why individuals' responded differently to the long and short set of questions.

If the tests fail in the context of the CPS, it is possible that another survey vehicle can be enhanced (in terms of its frequency of administration, or the nature of the labor force questions, for example) so an employment rate among adults with disabilities can be produced on a regular and timely basis. The evaluation of the underlying validity of the complete batteries of questions will permit a judgment as to which of these longer sets of questions should be used to generate employment data on persons with disabilities over time.

The Work Group and BLS will also consider the possibility of conducting an annual supplement to the CPS on disability.

## **Research questions to be addressed:**

There is agreement on the scientific validity of the proposed research, but there are several critical methodological issues to be addressed by the ERMM Work Group.

One important question involves the selection of the surveys to be used. Thus far, the list of candidates includes the Survey of Income and Program Participation (Census), the National Health Interview Survey (NCHS), and the Disability Assessment Schedule (WHO).

A second important question goes back to the definition of disability, and the survey thresholds used to operationalize the definitions. Once the definitions of disability are established, thresholds for disability must be set for the purposes of this research. Thresholds for disability need to be established in order to know who is and who is not included in the population. This is a particularly difficult task, but the organizations offering their survey for the tests will set thresholds for their own survey.

The last important question at this time involves establishing protocols for the research. Technical questions such as the appropriate sample composition and size, and the criteria for judging the acceptability of questions need to be addressed.

# **Members of the Employment Rate Measurement Methodology Work Group (Committee on Statistics)**

## ***Chair:***

*Philip Rones*, Assistant Commissioner  
Bureau of Labor Statistics, Office of Employment  
and Unemployment Statistics, Department of Labor

## ***Agency Members:***

*Access Board*

*Census Bureau*

### ***Department of Education:***

National Institute on Disability and Rehabilitation  
Research, Office of Special Education Programs

### ***Department of Health and Human Services:***

Centers for Disease Control and Prevention, National  
Institute on Mental Health, Office of the Assistant  
Secretary for Planning and Evaluation, Substance  
Abuse and Mental Health Administration,

### ***Department of Justice:***

Bureau of Justice Statistics

### ***Department of Labor:***

Bureau of Labor Statistics, Counsel for Labor  
Relations, Office of the Assistant Secretary for Policy

*Equal Employment Opportunity Commission*

*President's Committee on Employment of People  
with Disabilities*

*Social Security Administration*

# Appendix I:

## Executive Summary to the Annotated Bibliography

### Background

1. Executive Order 13078 requires that the Presidential Task Force on the Employment of Adults with Disabilities (PTFEAD) establish a work group on the measurement of the employment rate of adults with disabilities. The work group, chaired by the Assistant Commissioner for Employment and Unemployment Statistics of the Bureau of Labor Statistics (BLS), includes representatives of the major Federal data collection agencies, including BLS, the Census Bureau, and the National Center for Health Statistics, as well as agencies with a major research role in the area of disability statistics, including the Department of Education and the Social Security Administration. Other members of the work group are from other entities of the Department of Labor and the Department of Health and Human Services, as well as the President's Committee on the Employment of People with Disabilities, the National Council on Disability, the Equal Employment Opportunity Commission (EEOC) and the Department of Justice.
2. The purpose of this bibliography is to inform the work group of the PTFEAD as to what questions are available for identifying the adult population with disabilities and the degree to which these questions have been tested for reliability and validity. The work group proposed two alternative formats: a short battery of questions that could be used in a monthly survey and a longer set of questions that could be used on a supplement of a periodic survey. The short battery of questions could include a measure of severity and the longer version could provide a range of variables such as severity, duration, and type of disability.

### The ADA Definition of Disability

3. The Americans with Disabilities Act (ADA) of 1990 provides that "[n]o [employer] shall discriminate against a qualified person with a disability because of the disability of such individual."<sup>1</sup> A person with a disability is defined as one who has:
  - A physical or mental impairment that substantially limits one or more of the major life activities of the individual;
  - A record of such an impairment; or
  - Being regarded as having such an impairment.<sup>2</sup>
4. EEOC, the agency primarily responsible for enforcing the employment provisions of the ADA, has defined major life activities to include caring for oneself, performing manual tasks, walking, seeing, hearing, speaking, breathing, learning, sitting, standing, lifting, reaching, and working.<sup>3</sup> The major life activities under ADA appear to be at least as wide-ranging as what researchers in the field have labeled "Activities of Daily Living," "Instrumental Activities of Daily Living," and activities outside the home that are restricted because of barriers, including physical impediments and inadequate transportation.
5. EEOC has defined "substantial limits" as those restricting the duration, manner, or condition under which a person can perform a particular major life activity as compared to the average person in the general population.<sup>4</sup>

<sup>1</sup>42 U.S.C. 12112 (a)

<sup>2</sup>42 U.S.C. 12102 (2)

<sup>3</sup>29 C.F.R. 1630 (I); 29 C.F.R. pt. 1630, App. 1630.2 (I)

<sup>4</sup>29 C.F.R. 1630.2 (j)

## Collection of Employment Data and the ADA

6. While it is impossible to design a method to collect data on persons with disabilities that conforms directly with the ADA definition of disability, a data collection instrument can be designed that will generally rely upon the major elements of the ADA definition. The first such dimension is major life activities. This includes activities of a personal nature (breathing), of communication (seeing, hearing, speaking) and of participation (learning, working). Questions that address these types of activities have been used throughout the developed world by government statistical agencies and other researchers.
7. The second dimension of the ADA definition is the severity of the restriction experienced by an individual when he/she participates in those activities. Consideration has to be given to the nature of the restriction. Most disability surveys or disability modules address the issue of severity from the perspective of the individual only. Severity is assessed through questions such as able to do (a particular activity), partially able to do (a particular activity) or completely unable to do (a particular activity). While this is a good starting point, it is also important to understand the nature of the limitation. It may be that it is the environment within which the individual lives that causes the restriction. For example, the attitude of an employer toward a person with a disability may result in that individual reporting that he/she is completely unable to work; however, if the attitude of that employer changes, the barrier would be removed.

## Issues to be Considered when Deciding what Survey Vehicle to Use to Collect Data on Persons with Disabilities

8. The space available for disability-related questions and the type of interview to be used are two issues to be considered when choosing a survey vehicle to collect data on persons with disabilities.

## Available Space

9. The space available dictates the number of questions that can be asked. The census and multiple-topic surveys often restrict the amount of space that can be allocated to one particular topic. The Census, the American Community Survey (ACS), the American Housing Survey (AHS) and the Current Population Survey (CPS) are examples of multi-topic surveys that can accommodate a small number of questions on a particular topic. The 1991 Census of Jamaica included four questions on disability. The Kansas Department of Health and Environment's Program on Disability used the BFRSS and conducted a special disability survey using three questions to identify the population with disabilities.
10. Topic modules supplements offered by the Survey on Income and Program Participation (SIPP) provides the opportunity for a large number of questions to be directed to a particular topic.
11. A national survey or a census can be used as the sampling frame for a follow-up survey of disability. The National Health Interview Survey (NHIS) was used in this capacity for the 1994-95 NHIS-Disability survey. Canada used both their monthly Labor Force Survey and the Censuses of 1986 and 1991 to screen for persons with disabilities, and then conducted a follow-up survey on disability-related issues. The 1996 post-census survey in New Zealand was modeled after the 1986 and 1991 Canadian post-census surveys.
12. Both Australia and the United Kingdom designed surveys specifically for disability, developing the sampling frame from a list of addresses (UK) or a list of households (Australia). The 1998 NOD/Louis Harris Survey of Americans with Disabilities used a random sample of telephone numbers to generate the sample for their survey, as did the Indiana Independent Living Survey and the Idaho State Independent Living Council.



## Proxy versus Non-proxy Respondents

13. Many existing survey vehicles use proxy respondents to collect information about other household members. Some research has been done in Canada that shows that the use of proxy respondents reduces the overall disability rate. For the most part, the group most affected is those with minor disabilities. This may not be an issue for the PTFEAD initiative since the ADA definition is concerned with individuals who are substantially limited.

## International and National Experience with Defining Disability for Use in Surveys and Censuses

14. The two dimensions inherent in the ADA definition of disability have been the focus of many researchers in the area of information concerning persons with disabilities for the past three decades. Although not cited explicitly in this bibliography, the framework proposed by the World Health Organization (WHO) in their International Classification of Impairments, Disabilities, and Handicaps (ICIDH) has formed the basis for many of the data collection initiatives undertaken by national statistical agencies, State agencies, and other researchers.<sup>5</sup>
15. The Organization for Economic Cooperation and Development (OECD) formed a working group with representatives from Canada, Finland, France, West Germany, the Netherlands, Switzerland, the United Kingdom, and the United States to develop an approach to the measurement of the impact of ill health on daily living. Testing was undertaken by a number of the participating countries. The first four entries in the bibliography describe the development and the testing of the OECD activities of daily living questions.
16. Many developed countries have used a short battery of questions to collect information about persons with disabilities. The Census in the US, the Census in Jamaica, the NOD/Louis Harris Survey of Americans with Disabilities, the AHS, and the ACS have all used a short battery of questions. The BFRSS in Kansas also used three questions. Testing of these questions has been limited to focus groups.
17. The disability questions used in the NOD/Louis Harris Survey of Americans with Disabilities, the AHS, the ACS and the BFRSS in Kansas were designed to identify major life roles. No attempt was made to measure severity in these surveys.
18. The issue with the short list of questions is the ability to capture the severity of the disability. The broad method of questioning that is necessary to produce a short list of question has, to this point, precluded the development of a severity measure.
19. Canada conducted a test following the 1986 Census and the 1986 Health and Activity Limitation Survey (HALS) to determine the reliability of the 1986 Census question to measure the population with disabilities. The responses to the HALS screening questions (primarily OECD questions) — primarily obtained in non-proxy face-to-face interviews — were compared to the answers given to the Census questions — primarily respondent completed with no control for non-proxy. The result of the study showed that almost half of the population with disabilities as identified through the long list of HALS screening questions had indicated that they did not have a disability on the Census disability questions.

<sup>5</sup>The first version ICIDH — published in 1980 — provides a framework to examine the long-term consequences of disease, injury and disorders. While there have been many criticisms directed at this classification, the framework has proved to be a useful tool during the past two decades and many of the deficiencies and problems associated with it are currently being addressed in the revision process that has been underway since 1987. The revised version of the ICIDH is now being beta-tested in selected countries throughout the world.

20. A similar study is being undertaken with the revised disability question for the 2001 Census of Population in Canada and a sample of individuals who will be selected from the National Census Test. This sample will be asked the screening questions from the 1991 HALS survey. A micro match of the Census and HALS responses will be undertaken to determine if the new disability question will be more reliable in identifying the same population as the detailed screening questions from HALS. The results of this study will be available in the fall of 1999. At this point in time, Statistics Canada intends to conduct a post-census survey following the 2001 Census; there are no plans to release the disability data from the Census. The Census disability data will be made available only to users who specifically request that data. This is consistent with the position taken with respect to the 1986, 1991, and 1996 Census disability questions.

### **A Long List of Questions to Screen for Disability**

21. The OECD questions form the basis for the long list of questions that have been used in disability surveys conducted in the US, Canada, New Zealand, Australia, the Netherlands, Northern Ireland, and the United Kingdom. However, not all surveys used the same number of questions and often the wording of the question was changed slightly to reflect the culture of the country or the method of interviewing (proxy/non-proxy). Most countries have indicated that the extent of the testing involved holding focus groups with persons who were identified as having a disability and with other individuals involved with the population with disabilities (e.g., service providers, program managers, and academic researchers).
22. All of the surveys noted in the preceding paragraph have developed a severity scale based on the responses to the follow-up question that deals with able/partially able/completely unable. As these severity scales vary considerably, analysis will be required to select or develop a scale that is most appropriate for the purposes of the PTFEAD.
23. Not all individuals with a limitation in activity as a result of a psychiatric or mental health condition, a learning disability, or developmental delay identify using the OECD questions. The Canadian experience with a validation study conducted after the 1986 HALS indicated that additional questions were required to identify a larger proportion of these disability populations identified in the HALS entries as "mental" disabilities. These additional questions were added to the 1991 HALS.
24. Both Canada and New Zealand included questions about barriers to participation in various everyday activities. These questions were not asked as screening questions. Rather, they provided insight into the reasons why persons with disabilities were impeded or prevented from full participation in society. The "barriers" questions regarding employment were significantly developed between the 1986 and 1991 HALS, reflecting the focus of the 1991 HALS on employment equity issues.
25. Both the Idaho State Independent Living Council survey and the Indiana Survey of Independent Living used a longer list of questions to include the two dimensions of disability (i.e., major life activities, severity of disability). Some of the questions asked in the Indiana survey attempt to measure severity. The Lambeth Health Survey, used to estimate the prevalence of physical disability in a survey conducted in the 1960s in London, reports both validity and reliability testing.

## Recommendations

26. The 1996/99 Topical Module of SIPP appears to identify major life activities and severity of disability. More information is required concerning the type of testing undertaken to develop the questions. Some initial analyses of the wave 5 of the '96 panel could be undertaken and, if possible, comparisons to other data available from other national surveys should be made.
27. The Indiana Survey of Independent Living also provides a good definition of the two dimensions of disability, and the entire questionnaire has undergone some limited testing. This questionnaire is somewhat shorter than the SIPP questionnaire and, for that reason, may prove to be a useful alternative to SIPP. To facilitate further analysis, more information concerning the testing of the questionnaire should be obtained, as should a copy of the database. If possible, a comparison of the results to other Indiana State data available from national surveys should be undertaken.
28. The OECD questions have been used in many international and national surveys and they measure limitation in activity in a similar manner as defined by EEOC. These questions have been tested in a number of countries during the survey development phase.
29. Identifying persons with limitation in activity because of a psychiatric, learning, or developmental problem or condition requires additional questions. The 1991 HALS added some questions to include this population, and these questions were tested in focus groups and consultation meetings.
30. The WHO DAS II is currently being tested, and it has shown great promise in recent field trials. In contrast to other instruments, however, the WHO DAS II uses a 30-day recall period rather than 6 or 12 months.
31. The SF-36 and the SF-12 are two other instruments that should be examined in more detail as well as the results from the two field trials in North Carolina and Massachusetts.
32. Additional information should be obtained from EUROSTAT concerning their initiative to include an ad-hoc module on the European Union Labor Force Survey concerning the employment of persons with disabilities.
33. The measure of severity of disability is an issue that needs to be addressed if the measurement instrument developed by the work group of PTFEAD is to include this element in the screening questions. The severity scales that have been developed with the OECD questions and for the 1991 HALS (with the screening for mental disability questions) may not meet the needs of the PTFEAD.
34. Finally, questions concerning the environment asked in both the 1991 HALS and the 1996 New Zealand disability survey could be considered to measure the nature and extent of the barriers that prevent or impede persons with disabilities from securing and/or maintaining employment. The Craig Inventory of Environmental Factors (CHIEF) could also be considered.

## Appendix II:

### **A Research Plan for Evaluating Alternative Sets of Disability Screener Questions and Disability Assessment Surveys**

#### **Introduction**

This research plan represents one path (among many) toward the ultimate destination of designing and implementing a “statistically reliable and accurate method to measure the employment rate of adults with disabilities.” If the concept of disability were similar to other demographic concepts like age, gender, or educational attainment, the path to our destination would be relatively straight and narrow: We would design, evaluate, revise, and reevaluate a very brief set of questions for assessing disability. When we were satisfied with the reliability and accuracy of these questions, we would then introduce them into the demographic section of the Current Population Survey (CPS). But the concept of disability is anything but simple; in fact, it may be one of the most challenging concepts ever measured in a survey context.

What is it that makes the concept of disability so difficult to measure? First, from a legal perspective, the definition of disability (or more precisely, the interpretation of what a disability might be as specified by the Americans with Disabilities Act of 1991) is essentially unstable. While Congress has defined the term in the ADA, the courts determine key features of the definition of disability via case law. Second, from a theoretical perspective, the determination of whether a person is viewed as having a disability is very much context-dependent (i.e., disability is viewed as one possible outcome resulting from the interaction between a specific physical and/or psychological impairment and identifiable environmental factors). Significant changes either in the impairment or in the environmental context in which the person is situated (physical, social, cultural) may alter the way the person views himself/herself or the way in which others view the individual. Third, from a pragmatic perspective, impairments and contexts are myriad and, as a consequence, defy exhaustive meas-

urement via the survey method. These rather substantial obstacles notwithstanding, survey sponsors still endeavor to identify and categorize persons as having varying degrees of disabilities using questionnaires — though not all such survey instruments have explicit algorithms for this purpose. Moreover, these surveys differ in important ways as to the conditions under which data are collected (e.g., question wording, reference periods, concept definition, mode, method, sponsorship, etc.) and no one survey, as yet, has been identified on the basis of validation assessments as the “gold standard” for measuring disability. In sum, what we have available to us at this time is at least one reliable and accurate measure of employment (i.e., the CPS) and a variety of survey instruments that measure (independently or in combination) various physical and/or psychological impairments, activity limitations, and environmental factors. Some of these survey instruments have an algorithm for categorizing persons as disabled, while others do not. Most of these surveys differ — sometimes substantially — with respect to essential survey conditions. As a result, the estimates of disability they produce are not directly comparable.

Given the conditions described above, and assuming they provide a reasonably unbiased summary of the current State of disability measurement via the survey method, how might we proceed in our efforts to satisfy the mandate specified in Executive Order 13078? One set of research ideas for addressing this question is provided below.

#### **Proposed Research Design**

Ideally, as noted in the previous section, we would like to measure disability using a very brief set of quasi-demographic survey questions. But for all of the reasons mentioned previously, we are not optimistic about the feasibility of attaining this objective. Moreover, it may not even be possible to assess disability accurately and reliably using a longer, supplement-like “disability assessment survey” (or DAS). Recognizing that the collection of disability data via the survey method may not be optimal (but having no real alternative), what could we do in an effort to meet the mandate specified in Executive Order 13078?

*Overall Research Strategy.* The research strategy that we wish to propose would be to evaluate simultaneously multiple sets of disability screener-question sets (SQSs) and multiple disability assessment surveys (DASSs) in a split panel research design — each one of these instruments having the capacity (minimally) to categorize target persons as disabled or not disabled (see Figure 1). Our proposal would include four sets of screener questions and four DASSs (e.g., see Figure 2). The decision on which specific SQSs and which DASSs to include in the design would be negotiated within the framework of an organizational structure, but we believe there are several good candidates in each category. Moreover, it is not necessarily the case that these would have to be existing instruments; if so desired, the appropriate work groups could design and evaluate a new set of screener questions or a new DAS and test these new instruments against existing instru-

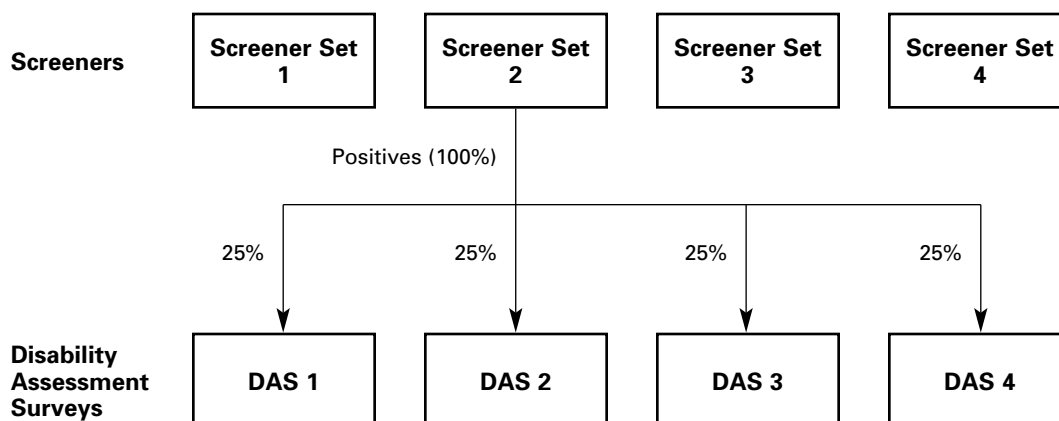
ments. We elaborate on various aspects of this research strategy below.

*Disability Screener Question Sets (SQSs).* At noted above, each disability SQS would need to be capable of categorizing target persons as disabled or not disabled (see Figure 1, cells A and D, respectively) — in other words, the instrument would need to have to be an explicit algorithm for doing so. It would also be desirable if each SQS had the capacity to distinguish among various general classes of disability on the basis of underlying impairment (i.e., physically-based or psychologically-based or both) and between cases on the basis of severity (i.e., severe or not severe). We recognize that such discriminations probably cannot be made using a limited set of screener questions; and even if attempted, we would expect substantial classification error. Moreover, if the CPS were to be consid-

**Figure 1. Basic Categorization Matrix**

		Condition/Disability	
		Yes	No
Test	Yes	Cell A: True Positive <i>[Sensitivity]</i>	Cell B: False Positive
	No	Cell C: False Negative	Cell D: True Negative <i>[Specificity]</i>

### Figure 2. Routing “Positive” Cases



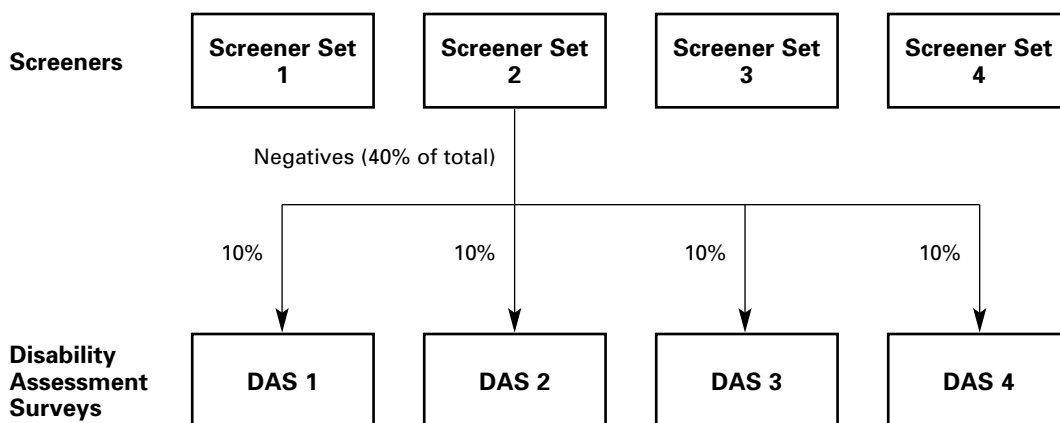
ered as a host survey for asking these screener questions, there would be restrictions on how much time could be allotted for this purpose (e.g., a maximum of three minutes a household)

*Disability Assessment Surveys (DASs).* Each DAS, ideally, would be capable of generating the same categorical data as the SQSs (i.e., via an explicit classification algorithm), but with greater accuracy and *with more detail* — for example, with respect to specific impairment categories (e.g., visual, auditory, motor, organic), existence of accomodative/obstructive environmental factors (e.g., physical, social, cultural), and amenability to medicinal interventions. Again, if the CPS were to be considered as a host survey for administering the DAS, there would be restrictions on how much time could be allotted for this purpose (e.g., a maximum of ten minutes a household). In addition, some effort would need to be made to minimize respondent burden in those cases where the target person is very severely disabled (e.g., psychoses, comatose States, quadriplegia).

*Essential Survey Conditions.* To satisfy Executive Order 13078, presumably one set of screener questions and/or one disability survey would need to be administered as part of the CPS (e.g., as control card items or as a periodic supplement, respectively). As a result, the essential survey conditions for this research should simulate the CPS survey context (e.g., CPS labor force questions, face-to-face and telephone interviewing, proxy responding, minimal respondent burden, BLS sponsorship).

*Design Specifics, Procedures and Analysis.* As noted above, we are proposing a rather elaborate split-panel research design that would incorporate four SQSs and four DASs. All of the instruments would need to be standardized in the sense that they would have to be capable of categorizing target persons as disabled or not disabled **at the time of the survey**; and the algorithms for doing so would have to be explicit — if the instrument designer/sponsor offers no such algorithm, one would need to be developed. The sequence of questioning would be as follows: CPS labor force questions (enough to categorize target persons as employed, unemployed, or not-in-the-labor-force) followed by one of the four SQSs followed by one of the four DASs. Individual interviewers would be trained to administer all four SQSs but only one of four DASs; they would be assigned a DAS on a random basis. If the research was conducted as part of the ongoing CPS, the test would be scheduled during a month when there was no supplement. As depicted in Figure 2, *all* target persons categorized as “disabled” by each SQS would be administered questions (via self or proxy reporting) from one of four DASs (i.e., the particular one the interviewer was trained to administer) — approximately 25 percent being allocated to each DAS. Unless evaluation data proves otherwise, we will assume that DASs are more accurate relative to the SQSs at identifying persons with disabilities. So, for example, if the SQS identifies the target person as disabled but the DAS does not, that person’s classification would constitute a *false positive*. As depicted by Figure 3, a

**Figure 3. Routing “Negative” Cases**



sample of persons categorized as “not disabled” would be allocated to each DAS; the goal here will be to check on *false negatives* (i.e., persons not identified as disabled by the SQS who *should* have been so classified). After the data collection phase has been completed, it should be possible to generate and review the following statistical data:

1. Basic classification data (e.g., disabled versus disabled) for all four SQSs
2. Basic classification data (e.g., disabled versus disabled) for all four DASs
3. Categorization matrices (see Figure 1) for all 16 SQS/DAS combinations (under the assumption that the DAS is the “true score”)

On the basis of these data alone, it should be possible to get a sense of which SQS is doing a “better” job classifying persons as disabled or not disabled. For example, we could compare each SQS/DAS “unit” against its “competitors” with regard to sensitivity (true positives), specificity (true negatives), false positives and false negatives. But defining “better” in terms of DAS-generated outcomes presupposes the validity of the DAS instruments. We need some independent measure of validity and/or data quality if we are to evaluate such findings.

*Validation and Survey Evaluation Research.* If resources were unlimited, in an effort to validate the survey data we had collected, we would conduct thorough post-administration assessments of each person’s physical and psychological status, activity limitations, and medical/medicinal regimens. We would also attempt to conduct an inventory of the environmental context (physical, social, and cultural) in which the individual was situated. However, we presume resources are limited and, so, we will need to consider other possible validation strategies. One feature of the CPS that might be useful as a partial validation measure is the opportunity that respondents have to volunteer information regarding the target person’s ability to work. To be more specific, when interviewers ask about work activity at several points in the CPS, there are response codes for answers such as “he [the target person] is disabled” or “I am unable to work”. We could conceivably use these data to validate responses to the SQS and the DAS. For example, if a person volunteered that the

target person was disabled, but neither the SQS or the DAS classified the individual as such, we’d have to question the sensitivity of the disability survey instruments. [Conversely, if the SQS or the DAS identified the person as disabled and the person was classified as employed by the CPS, we could learn something about accommodative and/or obstructive environmental factors by incorporating some follow-up probe questions to inquire about the existence of such factors.]

In addition to the partial validation measure described above, it should be possible to assess the accuracy of data being generated by the research instruments using standard survey evaluation techniques (e.g., cognitive interviews, respondent debriefing, behavior coding). For example, prior to fielding the survey, we could conduct a limited number of cognitive interviews on each SQS and each DAS in an effort to determine where the strengths and weaknesses of these instruments might be in terms of assessing impairments, environmental factors, and disability status. Then, on the basis of these cognitive interviews, we could design respondent debriefing questions — and ask them, when relevant, at the end of the household interview — that could directly probe respondents on the issue in question. For example, we may learn on the basis of cognitive interviewing that a person considers herself to have a disability (e.g., resulting from epileptic seizures) even though the medication she takes effectively controls those seizures. If a particular DAS does not specifically ask about medicinal regimens, we might want to include a probe question that addresses this issue. Evaluations of this sort will provide qualitative insights that would complement findings from the quantitative analyses described in the previous subsection.

Of course, a research effort of this scope and importance could not commence, let alone succeed, without substantial subject matter, methodological, and political support. Therefore, we would suggest that an organizational structure be developed early in the planning process that establishes this support network. We offer for review one such structure on the pages that follow.

## Organizational Structure and Group Membership

**Organizational Structure.** To succeed, the general research plan proposed above will require the collaborative efforts of an interdisciplinary team of project sponsors, subject-matter and policy specialists, survey methodologists, and operations specialists. The team would comprise seven groups with group-specific responsibilities (See Figure 4).

### 1. *Employment Rate and Measurement Methodology Group [Project Sponsors]*

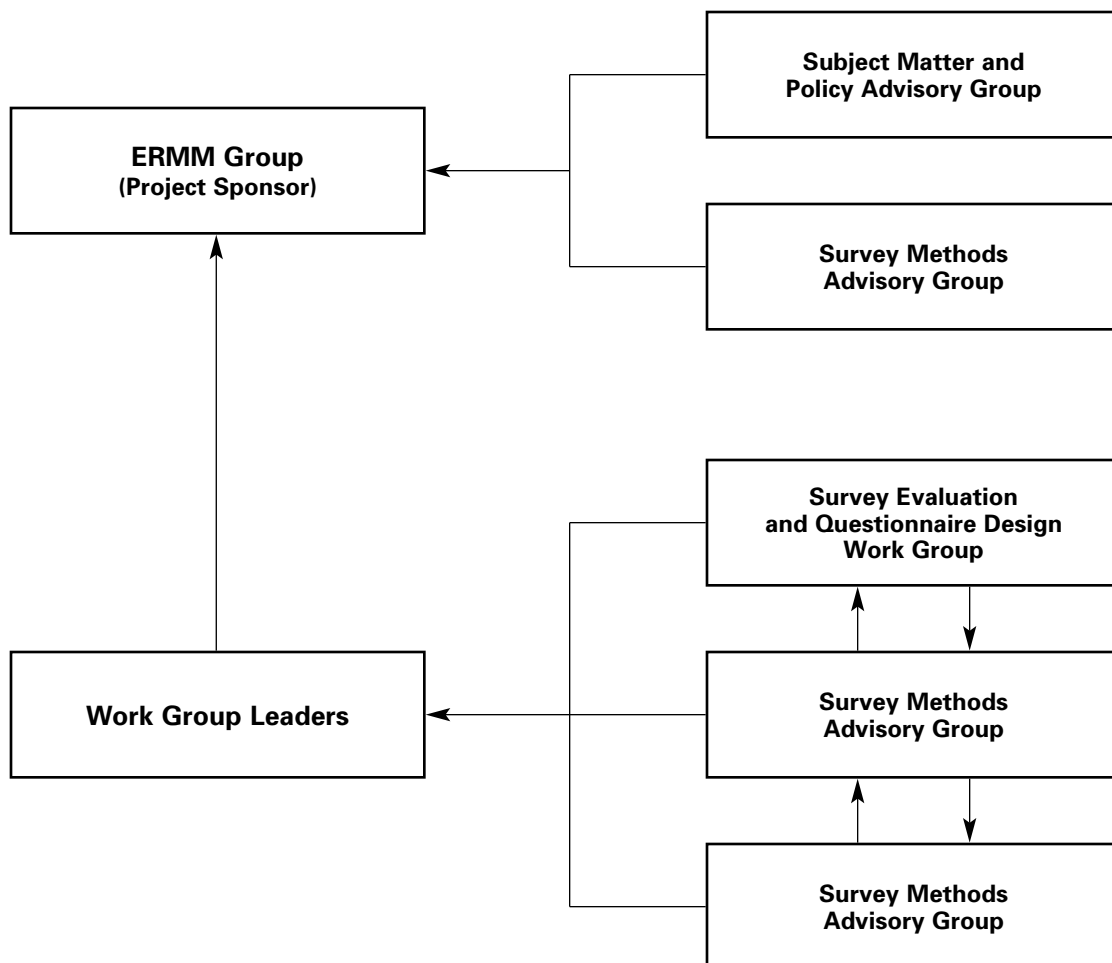
Responsibilities: To coordinate all efforts neces-

sary to design and implement a statistically reliable and accurate method to measure the employment rate of people with disabilities.

### 2. *Subject Matter and Policy Advisory Group*

Responsibilities: This group shall be responsible for providing the ERM group with subject matter and public policy expertise (as needed). This group will be asked to identify reports that summarize state-of-the-art research on disability issues. Included in such reports should be a listing of the various ways researchers have defined disability and a description of the most widely accepted framework(s) for conceptualizing dis-

**Figure 4. Schematic of Organizational Structure**



Note: In addition to identifying the reports described above, the advisory groups will be asked:

- to inform the ERM group as to how ADA definitions compare to those currently in use (with respect to similarities and differences);
- to suggest ways of resolving conceptual discrepancies, should they be found to exist;
- to suggest ways of operationalizing key ADA concepts/terms, if this has not already been done by other researchers; and
- to provide expertise (as needed) in developing a survey instrument that satisfies Presidential Executive Order 13078, section 2



ability (see **Note**). One or more policy experts will be asked to advise the ERMM group regarding the consistency of its decisions with respect to the Americans with Disabilities Act.

3. *Survey Methods Advisory Group*

Responsibilities: This group shall be responsible for providing the ERMM group with survey measurement expertise (as needed). This group will be asked to identify reports that identify and describe the various instruments/methods public and private research organizations have used to collect survey data on disability. Included here should be information on the actual survey questions used, survey documentation (e.g., question objectives), instructional materials (e.g., interviewer manuals), and survey evaluation reports (e.g., pretesting and/or quality assessment reports) — see **Note** at the bottom of the previous page.

4. *Work Group Leaders*

Responsibilities: To serve as liaisons between the ERMM group and their respective work groups, and to provide periodic updates as to the status of their efforts.

5. *Survey Evaluation and Questionnaire*

*Design Work Group*

Responsibilities: To evaluate existing survey questions and/or instruments that collect data on persons with disabilities, and to design and pretest a disability questionnaire that can be used within or as a supplement to the Current Population Survey (CPS). A formal evaluation of any set of questionnaire items would be expected to meet standards/guidelines set forth in *Pretesting Policy and Options: Demographic Surveys at the Census Bureau*, and in the BLS Commissioner's Order 2-96: *Ensuring Quality in the Data Collection Process*.

6. *Statistical Methods Work Group*

Responsibilities: To provide statistical methods support to the ERMM group and the other two work groups.

7. *Survey Operations Work Group*

Responsibilities: To perform all technical tasks associated with the development of a survey instrument for collecting data on persons with disabilities and to collaborate with the Survey Evaluation and Questionnaire Design work group in evaluating existing disability questions and/or survey instruments.

